

Identification of basins of attraction in the local Universe

Received: 26 February 2024

Accepted: 27 August 2024

Published online: 27 September 2024

 Check for updates

A. Valade^{1,2,3}✉, N. I. Libeskind¹, D. Pomarède^{1,4}, R. B. Tully^{1,5},
Y. Hoffman^{1,6}, S. Pfeifer¹ & E. Kourkchi^{1,5}

The structure in the Universe is believed to have evolved from quantum fluctuations seeded by inflation in the early Universe. These fluctuations lead to density perturbations that grow via gravitational instability into large cosmological structures. In the linear regime, the growth of a structure is directly coupled to the velocity field because perturbations are amplified by attracting (and accelerating) matter. Surveys of galaxy redshifts and distances allow one to infer the underlying density and velocity fields. Here, assuming the lambda cold dark matter standard model of cosmology and applying a Hamiltonian Monte Carlo algorithm to the grouped Cosmicflows-4 (CF4) compilation of 38,000 groups of galaxies, the large-scale structure of the Universe is reconstructed out to a redshift corresponding to $\sim 30,000 \text{ km s}^{-1}$. Our method provides a probabilistic assessment of the domains of gravitational potential minima: basins of attraction (BoA). Earlier Cosmicflows catalogues suggested that the Milky Way Galaxy was associated with a BoA called Laniakea. With the newer CF4 data, there is a slight probabilistic preference for Laniakea to be part of the much larger Shapley BoA. The largest BoA recovered from the CF4 data is associated with the Sloan Great Wall, with a volume within the sample of $15.5 \times 10^6 (h^{-1} \text{ Mpc})^3$, which is more than twice the size of the second largest Shapley BoA.

Coarse maps of the distribution of galaxies can be constructed by measuring redshifts. However, the redshift of a galaxy, z , is a combination of the cosmic expansion velocity and its peculiar motion, $cz = H_0d + v_{\text{pec}}$, where c is the velocity of light, d is the distance, H_0 is the Hubble constant (chosen to be consistent in the mean with distance measures) and v_{pec} denotes a galaxy's deviation from pure Hubble flow in the radial (line of sight) direction. Extensive surveys have been conducted over the last decades cataloguing the redshifts of millions of galaxies (for example, Center for Astrophysics (CfA)¹, Sloan Digital Sky Survey² and Six Degree Field³ surveys, among others). Such redshift surveys, using the approximation $d \simeq cz/H_0$, have uncovered an intricate network of clusters, filaments, sheets and voids—the so-called cosmic web. The regions of highest density delineated by rich clusters provided early

three-dimensional (3D) maps of the skeleton of the structure within $z = 0.1$ (refs. 4,5).

However, redshift maps give an incomplete picture. First, not all matter is luminous; most of the gravitationally active matter in the Universe is 'dark matter' with no electromagnetic manifestation⁶. In addition, galaxies are biased tracers of the density field⁷, sampling the density field in a nonlinear manner. The bias—the relation between galaxy number and matter density—is nonlinear and varies with the morphological type of the galaxy. Moreover, the counts of galaxies are sensitive to observational effects, such as the Malmquist bias.

On the other hand, galaxy peculiar velocities are unbiased tracers of the underlying matter density field because a galaxy's peculiar velocity (more aptly termed gravitational velocity) responds to the entire

¹Leibniz-Institut für Astrophysik Potsdam (AIP), Potsdam, Germany. ²Université de Lyon, Université Claude Bernard Lyon 1, CNRS/IN2P3, IP2I Lyon, Villeurbanne, France. ³Aix Marseille Université, CNRS/IN2P3, CPPM, Marseille, France. ⁴Institut de Recherche sur les Lois Fondamentales de l'Univers, CEA, Université Paris-Saclay, Gif-sur-Yvette, France. ⁵Institute for Astronomy, University of Hawaii, Honolulu, HI, USA. ⁶Racah Institute of Physics, Hebrew University, Jerusalem, Israel. ✉e-mail: avalade@aip.de

distribution of matter in the Universe, directly accounting for redshift distortions and independent of hidden mass biases. Hence, we opted to reconstruct large-scale structure from surveys of the galaxy distances and peculiar velocities.

The requirement for measuring a galaxy redshift is straightforward, but the measurement of distance is challenging. An extensive catalogue of distances (and hence peculiar velocities), such as the catalogue used here, is inevitably a heterogeneous combination of various surveys with varying sensitivities and selection functions. The Cosmicflows-4 (CF4) catalogue⁸ used here is the most extensive catalogue to date, comprising 56,000 galaxies. These are gathered in a composite catalogue of 38,000 constraints on the (linear) velocity field, composed of individual and groups of galaxies. The catalogue components are described in detail in the Methods. Succinctly, there is good all-sky coverage outside the zone of galactic obscuration to redshift $cz = 15,000 \text{ km s}^{-1}$, with an extension of good coverage in the galactic and celestial north to $cz = 30,000 \text{ km s}^{-1}$.

Because surveys of peculiar velocities of galaxies provide only their radial (line of sight) components, we are left with the problem of reconstructing the full 3D field. Under the assumption of potential flow, the 3D velocity and density fields can be inferred from such data (for example, the POTENT method⁹). However, peculiar velocities of galaxies are noisy, sparse and incomplete; consequently, the application of a POTENT-like method has a limited scope. A typical Cosmicflows data point has a relative error of about 20% in its measured distance, which corresponds to an uncertainty of $\sim 2,000 \text{ km s}^{-1}$ in the peculiar velocity for a galaxy at a redshift of $d \approx 100 \text{ h}^{-1} \text{ Mpc}$ ($cz \approx 10,000 \text{ km s}^{-1}$). For the lambda cold dark matter (Λ CDM) model¹⁰, this corresponds to a signal-to-noise ratio of about 15%, which is a very noisy signal. A Bayesian approach is needed to overcome such ‘bad’ data by incorporating the standard cosmological model as a prior. The Wiener filter and constrained realization algorithm^{9,11,12} provide an optimal linear reconstruction tool for such data. The Wiener filter/constrained realization method has been applied to reconstructions with peculiar velocity data^{13–16}.

Recently, forward, nonlinear Monte Carlo modelling^{17–21} in the context of Bayesian analysis has proven to be a successful mathematical tool for reconstructing realizations of the underlying density and velocity fields that are consistent with the input data and the assumed Bayesian prior model.

The Hamiltonian Monte Carlo Reconstruction of the Local Environment²² (HAMLET) method used in this study is one such approach. This method seeks to recover the Fourier modes of the density and velocity fields by assuming a cosmological model, Λ CDM¹⁰. The exploration of the parameter space is performed with a Hamiltonian Monte Carlo²³ (HMC) sampling algorithm and produces chains of states, in which each of these states represents a possible reconstruction of the velocity and density field consistent with the constraining data, their errors and the assumed Λ CDM model. The HAMLET method and its testing against mock data have been presented elsewhere^{22,24}.

In the following discussion, we focus on a particular feature of structures that we call a ‘basin of attraction’ (BoA). The entire Universe can be considered a patchwork of abutting BoA, just as the terrestrial landscape is separated into watersheds. A BoA is generally not gravitationally bound because the relative motion of distant points within it is usually dominated by cosmic expansion. The volume filling factor of BoA is unity; hence, their mean density is the cosmic mean. Streamlines diverge out of the local maxima of the velocity potential and converge onto its local minima, namely, they ‘stream’ away from the underdense to the dense regions of the Universe. In co-moving coordinates, streamlines (defined in the Methods) represent the trajectories along which matter flows. Any source position in a BoA leads through a streamline to a ‘sink’ near the potential minimum within the BoA. Thus, a BoA associated with a given local minimum of the potential is the volume that encompasses all streamlines that converge to a common region associated with a potential local minimum.

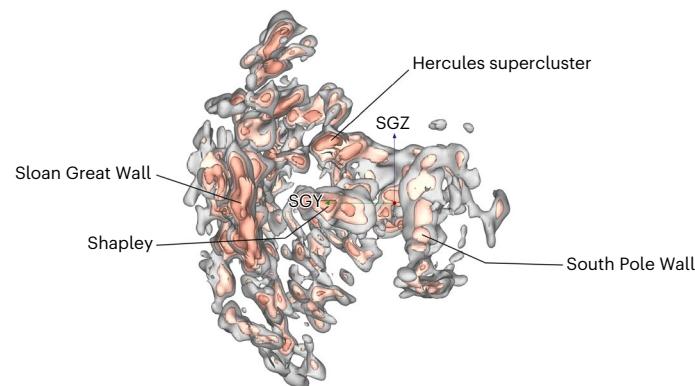


Fig. 1 | High-density perturbations δ in the distribution of matter in the local Universe with the mean of many HMC trials. The grey and red contours enclose regions of increasing density. Prominent structures appearing in this view are labelled. The coordinate system is supergalactic, and the Milky Way Galaxy is at the origin of the $10,000 \text{ km s}^{-1}$ long red (SGX), green (SGY) and blue (SGZ) axes. This view is looking inwards along the red axis from the negative SGX. Taken from Supplementary Video 1, frame at 2 min 8 s.

Although the definition of a BoA is precise if the distances and velocities are known without error, the data that inform their boundaries are noisy and incomplete. In previous studies^{25,26}, the BoA was computed on the mean inferred velocity field. This procedure is problematic because the construction of the BoA is not a linear process: the mean BoA from different realizations of the velocity field is not equal to the BoA built from the mean velocity field. An issue with previous BoA reconstructions based on the mean velocity field is that at the edge of the data, BoA may ‘infinitely leak’ outwards, as the smoothed velocity field converges to the null field. This deficiency calls for a more elaborate method, which is presented here.

Each HMC trial gives links between sources and sinks, consistent with the data. Because of imperfections in the data, the peculiar velocity fields and thus the basins of each state in the HMC chain differ from one realization to the next, while all are equally likely realizations of the velocity field constrained by the data and Λ CDM model. Thus, a BoA can be constructed for each state, resulting in an ensemble of BoA from which the probabilistic basin of attraction (p-BoA) is constructed. A detailed description of the construction of the p-BoA is given in the Methods. It is briefly noted that the core of the p-BoA is defined by the distribution of the sinks of the ensemble of BoA. The probability that a volume element is associated with a given p-BoA is given by the fraction of streamlines stemming from the volume element that terminates in the p-BoA. Our approach enables us to give two distinct probabilities: the probability that a BoA exists and the probability for each sampled point of space to belong to it. Given the probabilistic nature of the constructed BoA, some were robustly predicted with sharply defined boundaries, whereas others were less secure. Volume elements may be a part of several p-BoA, with probabilities summing to less than unity. As will be described, there is uncertainty in the boundary of the BoA that we live in.

Our results are presented by means of an accompanying video (Supplementary Video 1) that provides a visualization of the density and velocity structure of the volume surveyed by CF4, following our HMC analysis. The figures in this article are key frames of the video. The video begins with a display of the 3D distribution of galaxies/clusters sampled by CF4, transitioning to a vector representation of the measured radial peculiar velocities for the sample (blue, inwards; red, outwards). We then show the density and velocity field reconstructions from individual HMC trials before showing the averaged density field over many realizations (Fig. 1). The particularly prominent Sloan Great Wall, Shapley, Hercules and South Pole Wall structures were identified. More detailed features of high and low densities are identified in a slice

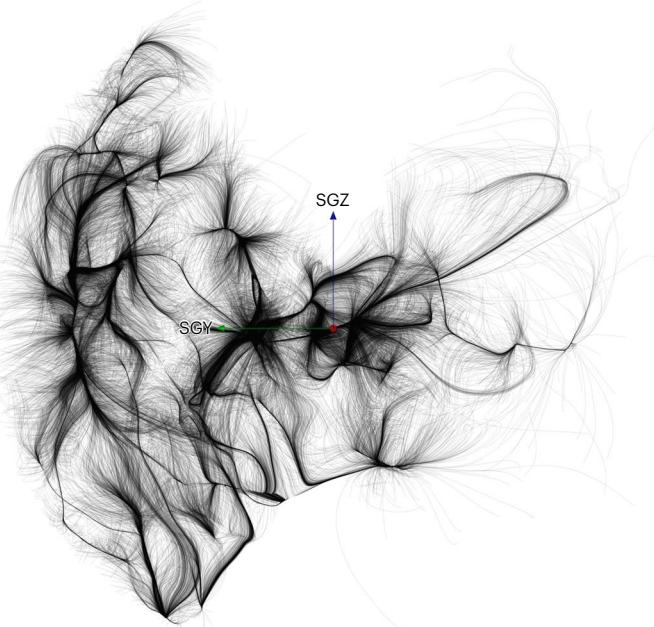


Fig. 2 | Streamlines constructed from the mean HMC velocity field. Looking inwards from the negative SGX with the Milky Way at the origin of the red, green and blue axes. Taken from Supplementary Video 1, frame at 3 min 34 s.

with thickness of $4,000 \text{ km s}^{-1}$ on the supergalactic equator ($\text{SGZ} = 0$), in which known elements of the cosmography, such as the 4-clusters or Sculptor Wall, are indicated^{27,28}.

The video then transitions to a display of the velocity streamlines from the mean reconstruction (Fig. 2). There are a multitude of sinks of the mean streamlines. The network of streamlines leading to each sink defines an individual BoA. Near the midpoint of the accompanying video, one sees the separate BoA defined by streamlines encased in shells of different colours (Fig. 3).

The reader is invited to explore the 3D large-scale structure in the accompanying interactive model, in which individual data points—galaxies and groups of galaxies—in the CF4 sample are shown as black points, an iso-surface in grey captures the reconstructed high-density regions and five particularly interesting BoA are encased (with distinctive colours) at surfaces given 50% probability from the chain of 1000 realizations. These five highlighted BoA include Shapley in yellow, a BoA centred in Ophiuchus associated with Laniakea in orange, Perseus–Pisces in cyan, Hercules in green and Sloan Great Wall in red.

In the latter part of the video, ambiguities in BoA boundaries are explored from the probabilities arising from individual realizations. The sink locations of scattered sources are shown, with their major aggregations identified with BoA (Figs. 4 and 5). A given source can be associated with a given BoA with a probability collected from many trials.

Streamlines sourced at the Milky Way find their way to two alternate sink regions. The most favoured terminus is in the region of the Shapley concentration^{29,30} (48% ending in a well-defined aggregate of sinks, enlarged to 58% with the inclusion of dispersed but proximate sinks). The remaining roughly 40% of the streamlines terminate in an aggregate in the vicinity of the Ophiuchus cluster. The associated BoA in this case very much resembles the Laniakea BoA identified using data from Cosmicflows-2 (ref. 25).

Within $cz = 15,000 \text{ km s}^{-1}$, where there is good all-sky coverage (outside the zone of obscuration), large-scale high-density structures found in earlier studies are recovered with generally improved acuity: the Shapley concentration at $cz \approx 15,000 \text{ km s}^{-1}$ (refs. 29,30), CfA Great Wall at $cz \approx 6,000 \text{ km s}^{-1}$ (ref. 1) and South Pole Wall at $cz \approx 11,000 \text{ km s}^{-1}$ (ref. 31), Perseus–Pisces filament at $cz \approx 6,000 \text{ km s}^{-1}$ (ref. 32) and

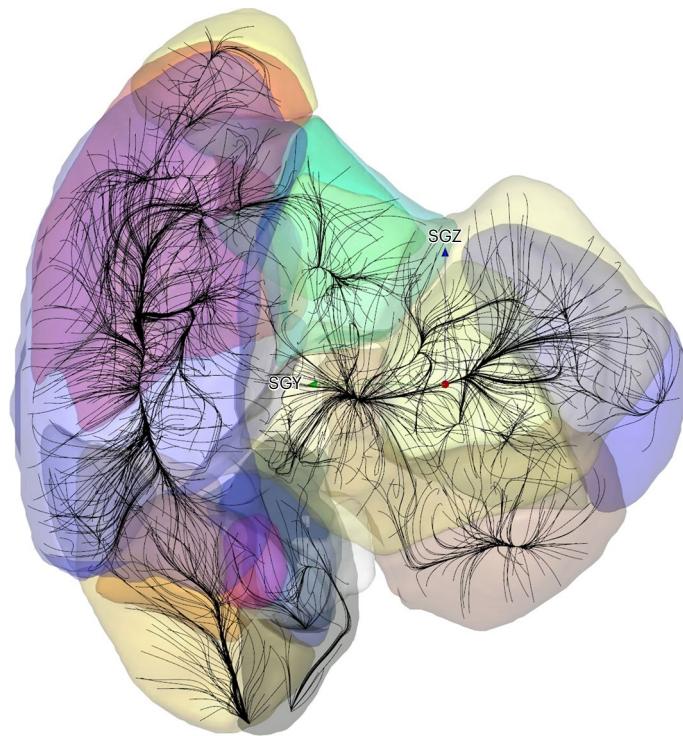


Fig. 3 | Velocity streamlines seeded at arbitrary locations within the reconstructed volume, with coloured envelopes associated with separate BoA extracted from the mean field. To counter the ‘infinite leaking effect’, these BoA have been cropped to the region covered by the data. Looking inwards from the negative SGX. Taken from Supplementary Video 1, frame at 4 min 18 s.

Hercules complex at $cz \approx 9,000 \text{ km s}^{-1}$ (ref. 33). Nearby, evidence emerges for a BoA centred in proximity to the highly obscured Ophiuchus cluster that lies behind the centre of the Milky Way Galaxy^{34,35} at $cz \approx 9,000 \text{ km s}^{-1}$. This BoA may include the so-called Great Attractor region³⁶ and the entity Laniakea²⁵, including ourselves. In the extension to $cz \approx 30,000 \text{ km s}^{-1}$, the Sloan Great Wall³⁷ and the associated structure are overwhelmingly dominant; indeed, it is the most prominent structure so far mapped in detail in our reconstruction.

Our methodology identified 37, 15 and 7 p-BoA with intrinsic probabilities of more than 20%, 50% and 75%, respectively. We named each basin after the over-density that hosted the aggregation of sinks from the HAMLET trials. For example, a plurality of sinks that lie within the high-density contours of the Shapley concentration defines the Shapley p-BoA. Extended Data Table 1 records probabilities for the realities of the most consequential nearby BoA, and Extended Data Table 2 provides the location of the cores and the associated volumes of the 15 p-BoA with probabilities greater than 50%, plus two of particular interest: Funnel and CfA Great Wall. The largest discovered p-BoA is associated with the Sloan Great Wall, encompassing a volume that is more than twice as large as the next largest one, Shapley.

Giving specific attention to the BoA called Laniakea from the analysis of Cosmicflows-2 data²⁵, two points are salient. First, this basin appears in only 62% of the realizations, with an attractor in the highly obscured region of the Ophiuchus cluster. Meanwhile, the Milky Way is contained by this basin only about 40% of the time. In fact, the Milky Way tends to be more often associated with the Shapley concentration (almost 50% of the time tightly and almost 60% of the time loosely). This leaves us in an interesting situation of supporting the likely existence of an Ophiuchus/Laniakea p-BoA with only ~40% probability of hosting our home galaxy. This option is explored in the illustration of Fig. 6. However, it is more probable that the Milky Way and a large volume of Laniakea are within the domain of the Shapley p-BoA.

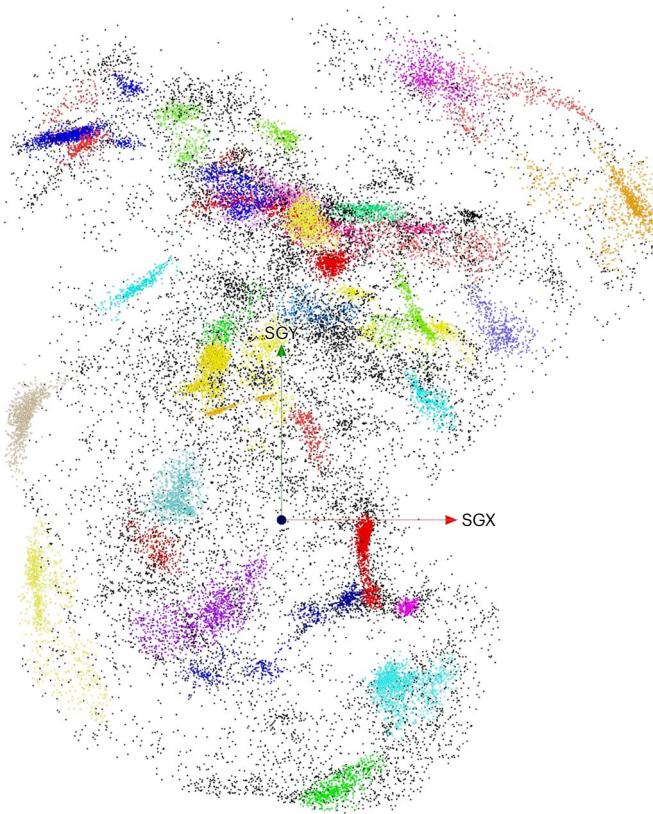


Fig. 4 | Sinks of velocity streamlines from individual HMC trials. Colours are associated with the major aggregates of streamline end points. Looking inwards from the positive SGZ. Taken from Supplementary Video 1, frame at 6 min 58 s.

There is a similar uncertainty with the status of the CfA Great Wall¹. The highest probability (~40%) linkage is to the Shapley BoA; alternatively, with probability ~1/3 is linked to the Hercules BoA; and alternately, with probability ~1/4 is a stand-alone CfA Great Wall BoA. This three-way ambiguity exists, even in a domain of dense data.

Given the finite depth of the Cosmicflows data and the linear increase in uncertainties with distance, the robustness of the reconstruction of the velocity field diminishes at the edge of the data. The dominant BoA identified here (Shapley, Hercules and the Sloan Great Wall and South Pole Wall) were all limited by the edges of our data. The current data are not deep enough to determine the outer bounds of these dominant p-BoA. It follows from the point of view of the BoA that cosmology has not yet reached its 'end of greatness'.

Methods

Data

Cosmicflows-4 is a collection of 56,000 galaxy distances gathered into 38,000 groups⁸. In total, there are eight distance measurement methodologies, although the great bulk comes from two methodologies involving correlations between global galactic photometric and kinematic properties: luminosity–rotation rate relation for spiral galaxies³⁸ (Tully–Fisher relation (TFR)) and luminosity–velocity dispersion–surface brightness relation for elliptical galaxies^{39,40} (fundamental plane (FP)). Most of the contributions for these two methodologies come from three sources. The largest contribution of 34,000 FP distances is confined to the sector north of the plane of the Milky Way and in the northern celestial hemisphere and extends to $z = 0.1$ in redshift⁴¹. The second FP contribution of 9000 distances is confined to the celestial south and redshift $z = 0.05$ (refs. 42,43). The third major source, that of 10,000 TFR distances⁴⁴, extends across the

sky, mostly within $z \approx 0.04$. The TFR material, dependent on kinematic information from radio telescopes, is particularly plentiful at celestial latitudes 0° to $+38^\circ$ accessed by the large Arecibo Telescope, deficient below latitude -45° accessed only by the smaller Parkes Telescope and given intermediate coverage otherwise with the Green Bank Telescope. The resultant coverage of the sky with an ensemble of contributions is uneven. Overall, there is reasonable coverage outside the zone of the Milky Way obscuration across the sky within $z = 0.05$ with a slight deficiency south of the Milky Way in the celestial north. There is a substantial extension evident in the figures north of the Milky Way and in the celestial north.

The scatter in distance and velocity measurements can be averaged over members of a galaxy group or cluster^{45,46}, thereby reducing errors. Each galaxy in CF4 is cross-matched to determine whether it lies within the region of collapse spatially and in velocity of a catalogued group. Groups can have from hundreds of distance measures (and independently, velocity measures) to just one. Most galaxies in CF4 stand alone in unassigned groups. The resulting catalogue contains 38,000 entries. Each entry consists of two angles defining positions on the sky and group-averaged redshifts and distance moduli with uncertainties. Unsurprisingly, uncertainties are lowest nearby and in rich groups, in which the density of information is greatest.

Linear theory

The complex statistics of the evolved universe can be reproduced only through elaborate simulations. For this study, we adopted the simpler modelling given by linear theory, which holds where deviations from the homogeneous and isotropic Friedmann model are small. In this context, the (over-)density field δ and the peculiar velocity field \mathbf{v} are bound by

$$-H_0 f(\Omega_m) \delta = \nabla \cdot \mathbf{v}. \quad (1)$$

This equation can be solved in either direction, that is, the density can be obtained from the velocity field and vice versa. The fields follow the statistics of multi-normal variables; thus, random realizations of the density field can be readily constructed from the matter power spectrum. In this study, we considered the one provided by the Planck mission¹⁰. In the Λ CDM standard model, the linear theory is generally considered an accurate description of the velocity field for scales above $\sim 5 \text{ Mpc } h^{-1}$ (where $h = H_0/100$). This sets the lower limit to the resolution of our reconstructions.

Bayesian model

The posterior probability was constructed as stated by the Bayes theorem applied to the reconstruction of the (linear) density field δ and the recovery of the distance of velocity tracers \mathcal{D} from coupled measurements of redshifts \mathcal{Z} and distance moduli \mathcal{M} :

$$P(\delta, \mathcal{D} | \mathcal{Z}, \mathcal{M}) = P(\mathcal{Z}, \mathcal{M} | \delta, \mathcal{D}) P(\delta, \mathcal{D}). \quad (2)$$

The details of these functions have been provided in a previously published paper²². Errors in the distance moduli were assumed to be normally distributed and are given in the CF4 catalogue. Errors in the redshifts were also assumed to be normally distributed with a fixed uncertainty of 50 km s^{-1} , augmented by the variance of the marginalized nonlinear part of the signal not recovered by the linear fields. We set the amplitude of the nonlinearities to $\sigma_{\text{NL}} = 100 \text{ km s}^{-1}$. This naturally handles log-normal bias^{16,47} and homogeneous Malmquist bias. The inhomogeneous Malmquist bias and selection effects were partially corrected and are the focus of ongoing studies. The model used in this study has been enhanced with respect to previous discussions^{22,24}: the selection function in the redshift present in the data has been modelled^{48,49}, and the computation of luminosity distances takes the peculiar movements of galaxies into account⁵⁰.

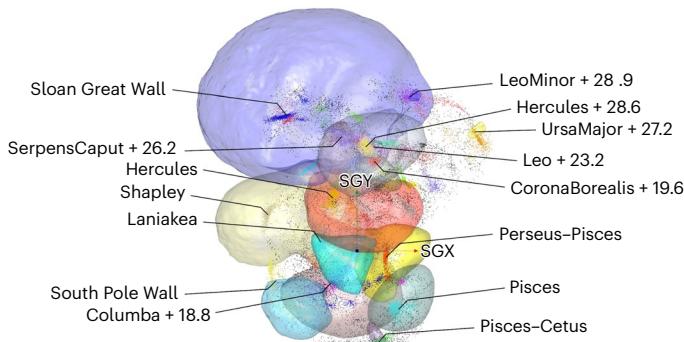


Fig. 5 | Envelopes of major BoA superimposed on the sinks of HMC trials shown in the previous figure. Names given to the major BoA. For previously unnamed objects, we used the convention ‘constellation + distance’ in units of $1,000 \text{ km s}^{-1}$. Note that several structures can be found in the same constellation (for example, Hercules). Taken from Supplementary Video 1, frame at 7 min 44 s.

Hamiltonian Monte Carlo

The parameter space was explored using the HMC method^{23,51}. The purpose of using a Monte Carlo chain was to produce a series of independent realizations that obey the probability law. This was achieved by an iterative process, in which each realization was constructed from the previous one. The variety of Monte Carlo methods were essentially derived from a variety of algorithms to perform this nuclear Monte Carlo step. The HMC method proposes a new realization by integrating the Hamilton equations of motion in a potential defined by $-\log P$ for a particle with an initial position that is the current realization and an initial momentum that is picked at random. This method has proven to be extremely efficient in high-dimensional spaces, in which other methods tend to stall. The meta-parameters of the exploration were tuned by the combined application of the no U-turn sampler⁵² and dual averaging step adaptation⁵².

Streamlines and basins of attraction

Streamlines are graphical visualizations of a velocity field. The equation of ‘motion’ of the line element of a given streamline $s(l)$, where l is the line parameter, is:

$$\frac{ds}{dl} = \mathbf{v}(l) \quad (3)$$

The choice of seeding points of streamlines is a matter of convention. Particles move along streamlines at a given moment; however, they traverse only a small fraction of a given streamline in the age of the Universe.

All cells with streamlines that converge to the same point in space, called the attractor, define a BoA. The study of BoA in a cosmological simulation⁵³ shows that these structures are stable in time, although a small basin may be absorbed by larger neighbours. Similarly, the adopted smoothing scale may lead to the merging of neighbouring basins.

Comment on terminology: in earlier studies^{25,54}, we gave the term ‘supercluster’ the same meaning we now give for BoA. Note that the term ‘watershed supercluster’ has also been used²⁶, as well as the related concept of ‘supercluster cocoons’⁵⁵. However, confusion arises because there is no community consensus on the meaning of the term ‘supercluster’. Because of the lack of an agreed upon definition, we avoid calling any structure a supercluster.

Probabilistic basins of attraction

The reconstruction of the flow field was done within a Bayesian framework, in which the HMC Markov Chain constructed states that sampled

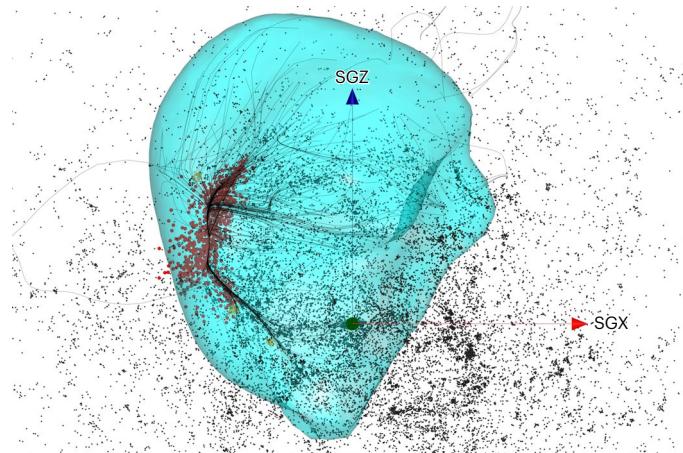


Fig. 6 | Cyan surface envelopes Ophiuchus/Laniakea BoA at 20% probability. Sinks from multiple trials shown as red points. Streamlines are from one of the realizations. Galaxies from the CF4 group sample are shown if $-9,500 < \text{SGY} < +4,500 \text{ km s}^{-1}$. Axes have lengths of $10,000 \text{ km s}^{-1}$. Taken from Supplementary Video 1, frame at 9 min 37 s.

the posterior probability distribution function, given the data and the assumed prior Λ CDM model. The HAMLET algorithm used in this study produced an ensemble of states from which an ensemble of BoA was constructed. The statistical distribution of the ensemble was consistent with the posterior probability. We extended the definition of a given BoA of a given state to a statistical BoA with a probability that reflects the distribution of the HMC states. We considered the ensemble of all sinks of streamlines of all states, excluding the sinks that lie outside of the volume covered by the data. We identified regions densely populated by sinks using the machine learning clustering algorithm HDBSCAN⁵⁶. This algorithm identifies aggregates of sinks and outperforms a simpler friends-of-friends method owing to its ability to select a density threshold per aggregate and the stability of its results with respect to its hyperparameters. Aggregates that were too small or too low in density were disregarded, whereas the remaining aggregates defined the core of the p-BoA. We define the intrinsic probability of a p-BoA based on the proportion of realizations in which it appears. The probability of a given source location on the grid belonging to a given p-BoA is defined by the proportion of realizations in which this source converges to a sink position within a given aggregate.

Data availability

The observational data used in this study were placed in the public domain in connection with the publication of Cosmicflows-4 (ref. 8). An interactive model illustrating the distribution of the Cosmicflows-4 sample of galaxies, an iso-contour of the mean HMC density reconstruction, mean velocity streamlines and the 50% probability inclusion shells of five prominent basins of attraction can be found at <http://sketchfab.com/3d-models/pboas-p05-cf4-mean-field-delta-and-velocity-9ba49209e60c48de8469b01ee8ee772e>. The estimated density and velocity fields are available upon reasonable request from the authors. We also offer an online tool that correlates distance and velocity using the constructed velocity field within this programme. Please access the tool at <http://edd.ifa.hawaii.edu/CF4calculator/>, and consult with the instruction on how to use the tool and interpret the results⁵⁷. We will continuously update the calculator for better visualization and incorporate the latest improvements.

Code availability

The code used to produce this study is not publicly available but can be communicated in response to reasonable requests.

References

1. de Lapparent, V., Geller, M. J. & Huchra, J. P. A slice of the Universe. *Astrophys. J. Lett.* **302**, L1 (1986).
2. York, D. G. et al. The Sloan Digital Sky Survey: technical summary. *Astron. J.* **120**, 1579–1587 (2000).
3. Jones, D. H. et al. The 6dF Galaxy Survey: final redshift release (DR3) and southern large-scale structures. *Mon. Not. R. Astron. Soc.* **399**, 683–698 (2009).
4. Tully, R. B., Scaramella, R., Vettolani, G. & Zamorani, G. Possible geometric patterns in 0.1c scale structure. *Astrophys. J.* **388**, 9 (1992).
5. Einasto, M., Einasto, J., Tago, E., Dalton, G. B. & Andernach, H. The structure of the Universe traced by rich clusters of galaxies. *Mon. Not. R. Astron. Soc.* **269**, 301–322 (1994).
6. Rubin, V. C., Ford Jr, W. K. & Thonnard, N. Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 (R=4kpc) to UGC 2885 (R=122kpc). *Astrophys. J.* **238**, 471–487 (1980).
7. Kaiser, N. Clustering in real space and in redshift space. *Mon. Not. R. Astron. Soc.* **227**, 1–21 (1987).
8. Tully, R. B. et al. Cosmicflows-4. *Astrophys. J.* **944**, 94 (2023).
9. Bertschinger, E. & Dekel, A. Recovering the full velocity and density fields from large-scale redshift-distance samples. *Astrophys. J. Lett.* **336**, L5 (1989).
10. Planck Collaboration et al. Planck 2015 results. XIII. Cosmological parameters. *Astron. Astrophys.* **594**, A13 (2016).
11. Hoffman, Y. & Ribak, E. Primordial Gaussian perturbation fields—constrained realizations. *Astrophys. J.* **384**, 448 (1992).
12. Zaroubi, S., Hoffman, Y., Fisher, K. B. & Lahav, O. Wiener reconstruction of the large-scale structure. *Astrophys. J.* **449**, 446 (1995).
13. Zaroubi, S., Hoffman, Y. & Dekel, A. Wiener reconstruction of large-scale structure from peculiar velocities. *Astrophys. J.* **520**, 413–425 (1999).
14. Doumler, T., Hoffman, Y., Courtois, H. & Gottlöber, S. Reconstructing cosmological initial conditions from galaxy peculiar velocities—I. reverse Zeldovich approximation. *Mon. Not. R. Astron. Soc.* **430**, 888–901 (2013).
15. Sorce, J. G. Minimization of biases in galaxy peculiar velocity catalogues. *Mon. Not. R. Astron. Soc.* **450**, 2644–2657 (2015).
16. Hoffman, Y., Nusser, A., Valade, A., Libeskind, N. I. & Tully, R. B. From Cosmicflows distance moduli to unbiased distances and peculiar velocities. *Mon. Not. R. Astron. Soc.* **505**, 3380–3392 (2021).
17. Kitaura, F. S. et al. Cosmic cartography of the large-scale structure with Sloan Digital Sky Survey data release 6. *Mon. Not. R. Astron. Soc.* **400**, 183–203 (2009).
18. Jasche, J. & Wandelt, B. D. Bayesian physical reconstruction of initial conditions from large-scale structure surveys. *Mon. Not. R. Astron. Soc.* **432**, 894–913 (2013).
19. Wang, H., Mo, H. J., Yang, X., Jing, Y. P. & Lin, W. P. ELUCID—exploring the local Universe with the reconstructed initial density field. I. Hamiltonian Markov Chain Monte Carlo method with particle mesh dynamics. *Astrophys. J.* **794**, 94 (2014).
20. Lavaux, G. Bayesian 3D velocity field reconstruction with Virbius. *Mon. Not. R. Astron. Soc.* **457**, 172–197 (2016).
21. Graziani, R. et al. The peculiar velocity field up to $z \sim 0.05$ by forward-modelling Cosmicflows-3 data. *Mon. Not. R. Astron. Soc.* **488**, 5438–5451 (2019).
22. Valade, A., Hoffman, Y., Libeskind, N. I. & Graziani, R. Hamiltonian Monte Carlo reconstruction from peculiar velocities. *Mon. Not. R. Astron. Soc.* **513**, 5148–5161 (2022).
23. Neal, R. MCMC using Hamiltonian dynamics. Preprint at <https://doi.org/10.48550/arXiv.1206.1901> (2011).
24. Valade, A., Libeskind, N. I., Hoffman, Y. & Pfeifer, S. Testing Bayesian reconstruction methods from peculiar velocities. *Mon. Not. R. Astron. Soc.* **519**, 2981–2994 (2023).
25. Tully, R. B., Courtois, H., Hoffman, Y. & Pomarède, D. The Laniakea supercluster of galaxies. *Nature* **513**, 71–73 (2014).
26. Dupuy, A. & Courtois, H. M. Dynamic cosmography of the local Universe: Laniakea and five more watershed superclusters. *Astron. Astrophys.* **678**, A176 (2023).
27. Courtois, H. M., Pomarède, D., Tully, R. B., Hoffman, Y. & Courtois, D. Cosmography of the local Universe. *Astron. J.* **146**, 69 (2013).
28. Pomarède, D., Hoffman, Y., Courtois, H. M. & Tully, R. B. The Cosmic V-Web. *Astrophys. J.* **845**, 55 (2017).
29. Raychaudhury, S. The distribution of galaxies in the direction of the ‘Great Attractor’. *Nature* **342**, 251–255 (1989).
30. Scaramella, R., Baiesi-Pillastrini, G., Chincarini, G., Vettolani, G. & Zamorani, G. A marked concentration of galaxy clusters: is this the origin of large-scale motions? *Nature* **338**, 562–564 (1989).
31. Pomarède, D. et al. Cosmicflows-3: the South Pole Wall. *Astrophys. J.* **897**, 133 (2020).
32. Haynes, M. P. & Giovanelli, R. in *Large-Scale Motions in the Universe: A Vatican Study Week* (eds Rubin, V. C. & Coyne, G. V.) 31–70 (Princeton University Press, 1988).
33. Einasto, M., Einasto, J., Tago, E., Müller, V. & Andernach, H. Optical and X-ray clusters as tracers of the supercluster-void network. I. Superclusters of Abell and X-ray clusters. *Astron. J.* **122**, 2222–2242 (2001).
34. Wakamatsu, K. I. & Malkan, M. A. Optical structure of the nearby cluster of galaxies in the direction of the galactic center: the optical counterpart of 4U 1708-23. *Publ. Astron. Soc. Jpn.* **33**, 57–66 (1981).
35. Hasegawa, T. et al. Large-scale structure of galaxies in the Ophiuchus region. *Mon. Not. R. Astron. Soc.* **316**, 326–344 (2000).
36. Dressler, A. et al. Spectroscopy and photometry of elliptical galaxies: a large-scale streaming motion in the local Universe. *Astrophys. J. Lett.* **313**, L37 (1987).
37. Gott III, J. R. et al. A map of the Universe. *Astrophys. J.* **624**, 463–484 (2005).
38. Tully, R. B. & Fisher, J. R. Reprint of 1977A&A....54..661T. A new method of determining distance to galaxies. *Astron. Astrophys.* **500**, 105–117 (1977).
39. Djorgovski, S. & Davis, M. Fundamental properties of elliptical galaxies. *Astrophys. J.* **313**, 59 (1987).
40. Dressler, A. et al. Spectroscopy and photometry of elliptical galaxies. I. New distance estimator. *Astrophys. J.* **313**, 42 (1987).
41. Howlett, C. et al. The Sloan Digital Sky Survey peculiar velocity catalogue. *Mon. Not. R. Astron. Soc.* **515**, 953–976 (2022).
42. Springob, C. M. et al. The 6dF Galaxy Survey: peculiar velocity field and cosmography. *Mon. Not. R. Astron. Soc.* **445**, 2677–2697 (2014).
43. Qin, F., Howlett, C., Staveley-Smith, L. & Hong, T. Bulk flow in the combined 2MTF and 6dFGSv surveys. *Mon. Not. R. Astron. Soc.* **477**, 5150–5166 (2018).
44. Kourkchi, E., Tully, R. B., Courtois, H. M., Dupuy, A. & Guinet, D. Cosmicflows-4: the baryonic Tully–Fisher relation providing 10 000 distances. *Mon. Not. R. Astron. Soc.* **511**, 6160–6178 (2022).
45. Tully, R. B. Galaxy groups: a 2MASS catalog. *Astron. J.* **149**, 171 (2015).
46. Tempel, E., Tuvikene, T., Kipper, R. & Libeskind, N. I. Merging groups and clusters of galaxies from the SDSS data. The catalogue of groups and potentially merging systems. *Astron. Astrophys.* **602**, A100 (2017).
47. Sorce, J. G., Stoica, R. S. & Tempel, E. Statistically bias-minimized peculiar velocity catalogs from Gibbs point processes and Bayesian inference. *Astron. Astrophys.* **679**, A1 (2023).

48. Strauss, M. A. & Willick, J. A. The density and peculiar velocity fields of nearby galaxies. *Phys. Rep.* **261**, 271–431 (1995).
49. Hinton, S. R., Kim, A. & Davis, T. M. Accounting for sample selection in Bayesian analyses. Preprint at <https://doi.org/10.48550/arXiv.1706.03856> (2017).
50. Calcino, J. & Davis, T. The need for accurate redshifts in supernova cosmology. *J. Cosmol. Astropart. Phys.* **2017**, 038 (2017).
51. Betancourt, M. A Conceptual introduction to Hamiltonian Monte Carlo. Preprint at <https://doi.org/10.48550/arXiv.1701.02434> (2017).
52. Hoffman, M. D. & Gelman, A. The no-u-turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. Preprint at <https://doi.org/10.48550/arXiv.1111.4246> (2011).
53. Dupuy, A., Courtois, H. M., Libeskind, N. I. & Guinet, D. Segmenting the Universe into dynamically coherent basins. *Mon. Not. R. Astron. Soc.* **493**, 3513–3520 (2020).
54. Pomarède, D., Tully, R. B., Hoffman, Y. & Courtois, H. M. The arrowhead mini-supercluster of galaxies. *Astrophys. J.* **812**, 17 (2015).
55. Einasto, M. et al. Multiscale cosmic web detachments, connectivity, and preprocessing in the supercluster SCI A2142 cocoon. *Astron. Astrophys.* **641**, A172 (2020).
56. Malzer, C. & Baum, M. A hybrid approach to hierarchical density-based cluster selection. *2020 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)* 223–228 (2020).
57. Kourkchi, E. et al. Cosmicflows-3: two distance–velocity calculators. *Astron. J.* **159**, 67 (2020).

Author contributions

A.V. developed the key ideas of this study, designed the reconstruction of the density and velocity fields, as well as the computation of the streamlines, and participated in the writing of this article. N.I.L., Y.H. and S.P. took active parts in the development of the methods described in this article, as well as in its redaction. D.P. analyzed the cosmography and generated all the visualizations (figures and videos)

that accompanied this study. R.B.T. led the writing of this article and was deeply involved in designing its scientific goals. E.K. built an online CF4 distance calculator.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s41550-024-02370-0>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41550-024-02370-0>.

Correspondence and requests for materials should be addressed to A. Valade.

Peer review information *Nature Astronomy* thanks Joseph Ribaudo and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2024

Extended Data Table 1 | Probability (in %) of some known features of the Local Universe to be part of probabilistic basins of attraction

	Ophiuchus BoA (Laniakea)	Shapley BoA	South Pole Wall BoA	Hercules BoA	CfA Great Wall BoA	Perseus BoA	Funnel BoA	Other
Milky Way	39 (38)	58 (48)	1	0	0	0	0	2 (12)
Shapley core (A3558)	0	99 (89)	0	0	0	0	0	1 (11)
Coma	0	40 (35)	0	33	25 (23)	0	0	2 (11)
South Pole Wall	0	21	42	0	0	0	0	35
Hercules supercluster	0	0	0	99	0	0	0	1
Perseus cluster	0	0	0	0	0	92	0	8
Funnel (A400)	0	0	0	0	0	72	6	22

Four claimed superclusters of the Sloan Great Wall³³ converge to the entity SCI 126 with a probability of about 95%. The percentages in brackets or alone given in this table are obtained by the automatic aggregation of attractor sinks by a machine learning algorithm. The larger values are based on a visual inspection in the 3D model of the locations of the scattered sinks with respect to the large aggregations of points.

Extended Data Table 2 | Properties of the p-BoAs appearing in Extended Data Table 1

Name	Prob %	Core position						Volume $10^6(h^{-1}\text{Mpc})^3$
		SGX $h^{-1}\text{Mpc}$	SGY $h^{-1}\text{Mpc}$	SGZ $h^{-1}\text{Mpc}$	R.A. deg	Dec deg	cz km/s	
Sloan Great Wall	99	-114.6 ± 15.0	221.0 ± 9.1	8.5 ± 22.0	195.9	-0.4	24909	15.51
Hercules	99	-39.6 ± 5.7	86.4 ± 8.6	79.5 ± 6.5	232.3	11.4	12391	1.86
Perseus cluster	93	47.8 ± 2.7	-18.3 ± 14.6	-32.7 ± 4.1	61.1	23.8	6074	1.06
Shapley	90	-145.1 ± 5.8	59.1 ± 10.7	-12.2 ± 8.7	201.6	-41.0	15715	7.02
Columba+13.8	89	-43.1 ± 18.6	-52.6 ± 12.8	-119.8 ± 3.5	88.0	-42.0	13774	2.69
Pisces	87	68.1 ± 11.5	-94.4 ± 8.4	20.3 ± 25.3	6.6	11.6	11817	0.66
Leo+23.2	76	27.2 ± 27.4	169.2 ± 8.4	-156.5 ± 11.7	141.9	13.4	23207	0.51
Hercules+28.6	72	14.7 ± 6.3	169.8 ± 7.3	229.3 ± 10.0	247.3	31.1	28569	4.11
South Pole Wall	66	-132.1 ± 11.5	-48.6 ± 26.7	18.7 ± 23.0	268.4	-65.9	14200	3.40
UrsaMajor+27.2	64	192.1 ± 18.3	180.7 ± 13.0	-67.0 ± 45.8	133.6	55.7	27210	0.48
Ophiuchus	62	-59.4 ± 7.0	14.6 ± 9.3	38.6 ± 16.1	245.3	-28.5	7233	0.80
CorBor+19.6	61	28.4 ± 5.1	145.7 ± 3.8	127.4 ± 7.1	233.6	38.4	19564	1.67
SerpensCaput+26.2	58	-25.1 ± 10.6	185.9 ± 8.4	183.5 ± 10.0	235.8	24.2	26237	2.19
Pisces Cetus	57	32.4 ± 11.0	-151.4 ± 7.3	-20.5 ± 12.6	17.3	-15.8	15621	0.81
LeoMinor+28.9	55	86.5 ± 12.1	250.9 ± 8.9	-113.5 ± 14.5	153.0	31.6	28868	0.54
Funnel	24	71.7 ± 3.2	-49.6 ± 2.6	-37.5 ± 3.7	45.6	18.7	9491	1.02
CfA Great Wall	23	-28.4 ± 12.7	64.1 ± 3.9	11.9 ± 10.1	202.4	5.0	7111	0.61

Each core position is the mean of the ensemble of sinks from realizations that constitutes the p-BoA. The associated volume is the mean of the volumes of each realization of the p-BoA and the error bars represent the scatter over the ensemble of states. The p-BoAs are sorted by their statistical significance.